H-EQPE Model and L-Checkpoint Algorithm: 
A Decision-Guidance Approach for Detecting Hypoglycemia of Diabetes Patients

Chun-Kit Ngan, Division of Engineering and Information Science, The Pennsylvania State University, Malvern, PA, USA

Lin Li, Department of Computer Science and Software Engineering, Seattle University, Seattle, WA, USA

ABSTRACT

The authors propose a Hypoglycemic Expert Query Parametric Estimation (H-EQPE) model and a Linear Checkpoint (L-Checkpoint) algorithm to detect hypoglycemia of diabetes patients. The proposed approach combines the strengths of both domain-knowledge-based and machine-learning-based approaches to learn the optimal decision parameter over time series for monitoring the symptoms, in which the objective function (i.e., the maximal number of detections of hypoglycemia) is dependent on the optimal time point from which the parameter is learned. To evaluate the approach, the authors conducted an experiment on a dataset from the Diabetes Research in Children Network group. The L-Checkpoint algorithm learned the optimal monitoring decision parameter, 99 mg/dL, and achieved the maximal number of detections of hypoglycemic symptoms. The experiment shows that the proposed approach produces the results that are superior to those of the domain-knowledge-based and the machine-learning-based approaches, resulting in a 99.2% accuracy, 100% sensitivity, and 98.8% specificity.

Keywords: Decision Support, Diabetes, Hypoglycemia, Optimization Model, Parameter Learning, Time Series

1. INTRODUCTION

Diabetes, the eighth leading cause of death in the world (World Health Organization, 2014), is due to either the body not producing enough insulin or the body not using insulin properly, which causes blood glucose levels (BGLs) to rise higher than normal. This high BGL leads to
common symptoms of diabetes such as urinating often, feeling very thirsty, increasing hunger, and causing many complications (e.g., heart disease, stroke, kidney failure, etc.).

According to the statistical data from the 2014 National Diabetes Statistics Report (Centers for Disease Control and Prevention, 2014), 29.1 million Americans (i.e., 9.3% of the U.S. population) had diabetes in 2012, and 25.8 million Americans (i.e., 8.3% of the U.S. population) had diabetes in 2010 (American Diabetes Association, 2014). This 12.79% increase of diabetes patients just within two years has drawn a significant attention from the medical domain. To maintain and control a normal BGL, diabetes patients can be treated by injecting insulin. The major effect of injecting insulin on diabetes patients is to lower their BGLs to normal, but it might cause an insulin reaction, often referred to as hypoglycemia, that is, a condition characterized by an abnormally low BGL. Hypoglycemia can cause serious consequences, such as fatigue, dizziness, confusion, headaches, blurred vision, seizures, unconsciousness, or even death, on the patients. In order to detect whether a patient is experiencing hypoglycemia, doctors typically check the patient’s BGL routinely. A big challenge is how to accurately and effectively determine the BGL monitoring threshold to prevent the patients from suffering hypoglycemic events. This is exactly the focus of this paper.

Currently, the existing approaches to identify and detect hypoglycemia based on the routine BGL check can be roughly divided into two categories: domain-knowledge-based and machine-learning-based. The former approach relies solely on medical experts’ knowledge. Based upon their knowledge and experiences, medical experts have identified a BGL monitoring threshold, i.e., 70 mg/dL (American Diabetes Association, 2014), which can be used to detect the occurrence of hypoglycemic events. When the diabetes patients’ BGLs become lower than 70 mg/dL, an alarm that signals the patients suffering the symptoms of hypoglycemia is triggered. Often this threshold may work well with a number of patients since it is determined by the medical experts based on their past experiences, observations, intuitions, and domain knowledge. However, it is not always accurate, as different patients have different health conditions; hence one value does not fit them all. In addition, the threshold is static, but the problem that this paper deals with is often dynamic in nature. Each patient’s health condition is constantly impacted by many unknown and uncontrollable factors and the surrounding environment, e.g., temperature, food, weather, etc. The domain-knowledge-based approach lacks a formal mathematical computation that dynamically learns the monitoring parameter to meet the needs of each patient’s health condition.

An alternative approach is to utilize machine-learning-based methods such as logistic regression models (Bierens, 2008; Cook, et. al., 2000; Dougherty, 2007; Hansen, 2010; Heij, et. al., 2004; Rawlings, et. al., 2012; SL, et. al., 2014) and Bayesian classifications (North, 2012; Provost & Fawcett, 2013; Tan, 2014; Zaki, et. al., 2014). The logistic regression (LR) model is a discriminative model that predicts the occurrence of hypoglycemic symptoms (i.e., 0 means no symptom; 1 means the occurrence of a symptom) by learning regression coefficients of the logistic distribution function of the explanatory variables. The Bayesian classifier (BC) uses the Bayes theorem to predict the occurrence of hypoglycemia as the one that maximizes the posterior probability. Both the regression coefficients and the posterior probability are learned based on the historical data by applying nonlinear regression models and maximum likelihood estimation (MLE) (Myung, 2003). The LR model building and the BC construction both are an iterative process, including model/classifier formulation, parameter/probability estimation, and model/classifier evaluation, which makes learning parameters/probability by mathematical computations complicated and computationally expensive. Despite enormous improvements in computer software in recent years, fitting such nonlinear quantitative decision model (Evans, 2010) is not a trivial task, especially if the parameter learning process involves multiple explanatory variables, i.e., high dimensionality. Working with high-dimensional data creates a difficult challenge, i.e., a
Related Content

Principal-Agency Relations in Organizational Networks
[www.igi-global.com/article/principal-agency-relations-in-organizational-networks/170606?camid=4v1a](www.igi-global.com/article/principal-agency-relations-in-organizational-networks/170606?camid=4v1a)
A Study of Information Requirement Determination Process of an Executive Information System
www.igi-global.com/chapter/study-information-requirement-determination-process/11324?camid=4v1a

Classification and Ranking Belief Simplex
www.igi-global.com/chapter/classification-ranking-belief-simplex/11242?camid=4v1a

Data Mining in Gene Expression Analysis: A Survey
www.igi-global.com/chapter/data-mining-gene-expression-analysis/28158?camid=4v1a